Presentation Attack Detection (PAD)

1. Overview

Aadhaar's integration of facial authentication across assisted and remote scenarios makes it a valuable, contactless verification method for millions of citizens. However, it also introduces vulnerabilities to a range of Presentation Attacks (PAs), from simple printed photographs to highly sophisticated deepfakes and adversarial AI manipulations. While UIDAI has baseline PAD measures in place, these defenses must be systematically evaluated against contemporary attack methods, and vulnerabilities must be addressed through advanced, field-ready solutions. This approach of comprehensive evaluation and targeted solution development shall enable UIDAI to strengthen Aadhaar's authentication framework in a proactive, evidence-based manner.

2. Objectives

1. **Discover Vulnerabilities**: To explore and evaluate the effectiveness of the Presentation Attack Detection (PAD) algorithms currently employed by the UIDAI team against attacks such as:

    a) Print Attacks - Using printed photographs of a person's face to trick the facial recognition system into a false match

    b) Replay Attacks - Displaying pre-recorded videos or images of a person's face on a screen to spoof authentication

    c) Mask-based Attacks - Wearing 2D or 3D masks (including silicone or resin masks) to impersonate a target's facial features

    d) Digital Manipulation Attacks (including retouching and morphing) - Altering genuine images through editing (e.g., airbrushing, morphing two identities) to bypass face-matching safeguards

    e) Deepfake Attacks - Using AI-generated synthetic media to realistically mimic a target's face and expressions in video or image form

    f) Adversarial Attacks -Crafting subtle, machine-targeted perturbations in images or video frames to deliberately mislead AI-based face recognition or PAD algorithms.

    g) Source Detection – Identify the solution to ascertain that the source of the image/video is from an authentic device/camera and has not been tampered with.

2. **Fix Vulnerabilities:** To provide a holistic, integrable solution in the form of AI/ML models and related techniques to address the vulnerabilities uncovered in the current algorithms, ensuring improved resilience against known and emerging presentation attacks.

3. Challenge Description

UIDAI is seeking proposals with a mechanism and a solution that:

(a) rigorously evaluate the effectiveness of current Aadhaar PAD algorithms against the specified attack categories, and

(b) develop and deliver enhanced PAD models that can be integrated into Aadhaar's authentication framework.

The solution should be capable of real-time or near-real-time detection, function in varied field conditions (low-light, low-connectivity, device variability), and maintain a balance between detection sensitivity and user convenience. Solutions should also include the creation of PAD datasets representative of India's diverse demographics and usage conditions, along with actionable recommendations for continuously enhancing the algorithms.

4. Expected Key Results

a) Conduct systematic testing of existing Aadhaar PAD algorithms against the specified attack categories and documenting the results

b) Develop AI/ML-based PAD models capable of mitigating vulnerabilities uncovered in the evaluation phase

c) Support real-time or near-real-time detection with minimal computational overhead

d) Ensure compatibility with UIDAI's authentication APIs and standards, device certification processes, and operational workflows

e) Demonstrate robustness across diverse environmental conditions, demographics, and hardware categories

f) Include optional creation or augmentation of annotated PAD datasets representing Indian demographics and realistic attack scenarios

g) Provide a configurable sensitivity/risk scoring system and detailed technical documentation for integration

5. Evaluation Criteria

a) Evaluation Depth: Rigor of analysis of the current Aadhaar PAD algorithm's performance against the specified attack categories

b) Accuracy & Effectiveness: Improvement in detection performance compared to baseline UIDAI algorithms

c) Robustness: Consistency of performance across varied environmental, demographic, and device conditions

d) Integration Feasibility: Ease of deploying the new PAD solution into Aadhaar's existing authentication infrastructure

e) Innovation: Novelty of the detection methods, adaptability to emerging attack vectors, and inclusion of adversarial resilience techniques

f) Compliance & Privacy: Adherence to UIDAI's privacy-by-design principles, secure biometric handling, and compliance with Aadhaar Act provisions

g) Scalability: Ability to handle Aadhaar-scale authentication volumes without significant latency or resource strain